Reprinted from JOURNAL OF APPLIED PHYSICS, Vol. 39, No. 10, 4828–4836, September 1968

Hertzian Fracture in Single Crystals with the Diamond Structure

B. R. LAWN
School of Physics, University of New South Wales, Kensington, New South Wales, Australia
(Received 15 May 1968)

Extension of an earlier theory of Hertzian fracture in brittle isotropic materials is here made to include the case of brittle single crystals, with particular reference to crystals having the diamond structure. A detailed description is first given of the inhomogeneous stress field in a flat, elastic specimen loaded normally with a hard sphere. The geometry of cracks growing in such a stress field is then discussed, taking into account the anisotropy in surface energy relevant to the diamond-structure crystals. Analyses of the mechanics of crack growth into the crystals subsequently indicate that for a certain range of indenter size, the Hertzian crack passes through four equilibrium stages, as it does in glass, before reaching its fully developed length. As a result Auerbach's law, which states that the critical load on a spherical indenter necessary to produce a fully developed Hertzian fracture is proportional to the radius of the indenter, holds within this certain range of indenter size. This law is experimentally confirmed by Hertzian fracture tests on single crystals of silicon. The mechanism of crack initiation and growth outlined in this paper is then discussed in terms of conclusions made by previous workers on the nature of the Hertzian cracks in the diamond structure crystals. Finally, possible application of the Hertzian test to the study of some mechanical properties related to the fracture surface energy in brittle single crystals is indicated.

INTRODUCTION

In recent papers, Frank and Lawn[1,2] treated the geometry and mechanics of Hertzian fracture in brittle, isotropic materials in detail. In particular, a formal derivation of the long-established empirical Auerbach law[3] was presented. This law states that the critical load P_c required for a hard, spherical indenter to produce a cone crack in a flat, brittle specimen is proportional to the radius r of the indenter. Particular interest stems from the fact that the theoretical analysis predicts that the constant of proportionality is linearly dependent on γ, the fracture surface energy which, for ideally brittle materials, is simply the free energy required to create unit area of surface. Furthermore, this proportionality constant is independent of the size of the inherent flaws from which the Hertzian cracks initiate, so that Auerbach's law affords, in principle, a simple and unique means of tabulating information about γ without resort to any knowledge of flaw statistics.

The investigation of Hertzian fracture mechanisms in single crystals with anisotropic surface energies and elastic moduli is of interest for the following reasons. First, because of recent widespread activity in measurements of the γ term in single crystals, it is desirable to establish the conditions, if any, under which Auerbach's law may be applicable. It might be argued that the anisotropic nature of a single-crystal specimen could possibly modify the fracture process to such an extent that Auerbach's law ceases to hold; it was, for example, predicted that the introduction of a sliding motion to a loaded indenter causes a sufficiently large perturbing effect on the stress field in the specimen, even for very low coefficients of friction between indenter and specimen, to destroy the Auerbach behavior.[2] That crystallographic anisotropy does play a significant role in Hertzian fracture may be readily deduced from the observations of Howes and Tolansky on diamond (for a review see Ref. 4): the fracture geometry and critical fracture load are both very much dependent on the crystallographic orientation of the diamond surface. Second, the mechanism of Hertzian crack formation in the diamond-type crystals has proved to be a subject of considerable interest in its own right; e.g., the interpretation of certain features associated with the Hertzian cracks in terms of dislocation processes accompanying fracture has formed the basis of some controversy. We return to this point in the discussion.

It is intended that much of the treatment which follows be applicable to brittle single crystals in general. However, crystals possessing the diamond structure are cited as particular examples for study for two reasons; (i) They are highly brittle at room temperatures and γ is therefore presumably determined principally by the reversible surface energy. The variation of γ with crystallographic orientation is sufficiently well known in these materials, and appears to be far more significant in accounting for cleavage tendencies than the anisotropy in the elastic constant.[5] This is a matter of some convenience, for it indicates that we can, to a first approximation, consider our crystals to have anisotropy in γ only; for the purpose of describing the stress distribution under a spherical indenter, the assumption of elastic isotropy leads to great mathematical simplification; (ii) Because of the demand for high-quality semiconductor materials, large, nearly perfect single crystals of silicon and germanium are readily obtainable commercially. With such materials, complications in the crack growth due to lattice imperfection and microstructure do not arise.

[1] F. C. Frank and B. R. Lawn, Proc. Roy. Soc. **A299**, 291 (1967).
[2] B. R. Lawn, Proc. Roy. Soc. **A299**, 307 (1967).
[3] F. Auerbach, Ann. Phys. Chem. **43**, 61 (1891).
[4] V. R. Howes, *Physical Properties of Diamond*, R. Berman, Ed. (Clarendon Press, Oxford, 1965), Chap. 6.
[5] G. A. Wolff and J. D. Broder, Acta Cryst. **12**, 313 (1959).

The Hertzian Stress Field

We now consider the elastic stress situation in a flat-surfaced specimen beneath a normally loaded spherical indenter prior to fracture occurring. Using the same notation as in the paper by Frank and Lawn,[1] the radius a of the circle of contact between indenter and specimen is given by

$$a^3=(4/3)(k/E)Pr, \tag{1}$$

where E is the Young's modulus of the specimen, P is the load, r is the radius of the ball, and k is a dimensionless constant involving the Young's modulus E, E', and Poisson's ratio ν, ν', of specimen and ball, respectively. The factor k is written

$$k=(9/16)[(1-\nu^2)+(1-\nu'^2)E/E']; \tag{2}$$

it has value unity when indenter and specimen are of the same material, and reduces to one-half for a rigid indenter. The mean pressure,

$$p_0=P/\pi a^2, \tag{3}$$

serves as a convenient unit of stress.

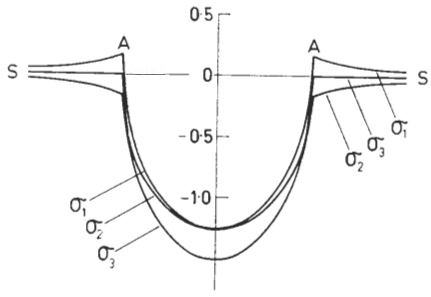


FIG. 1. Variation of principal stresses in the surface of a semi-infinite elastic medium (SS) in contact with a spherical indenter. The stresses are measured along a line coincident with a diameter (AA) of the contact circle. p_0 is the unit of stress.

The Hertzian stress field in an isotropic specimen can be readily calculated from equations given in papers by Huber[6] and Hamilton and Goodman.[7] For the purpose of discussing probable crack paths in single-crystal specimens it is necessary to describe this stress field in a little more detail than before.[1] We investigate the *distribution* of the three principal stresses, σ_1, σ_2, σ_3, beneath the indenter, in Figs. 1 and 2, and their *directions*, by means of stress trajectories, in Fig. 3. All stresses, expressed in units of p_0 in these figures, are calculated on the basis that $\nu=\frac{1}{3}$. AA indicates the region of contact and AS the region of free surface in all cases. The principal stresses, σ_1, σ_2, σ_3, are so labeled that $\sigma_1>\sigma_2>\sigma_3$ nearly everywhere; σ_3 *does* exceed σ_2 in a shallow region below the free surface AS in Fig. 2. We now proceed

[6] M. T. Huber, Ann. Physik **14**, 153 (1904).
[7] G. M. Hamilton and L. E. Goodman, J. Appl. Mech. **33**, 371 (1966).

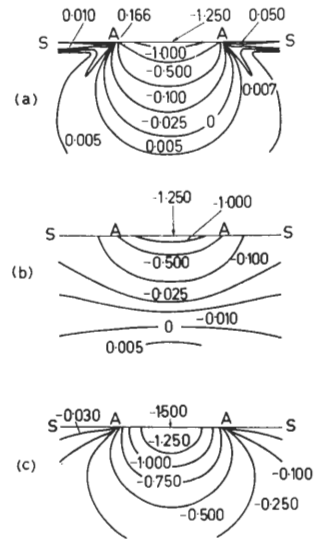


FIG. 2. Contours of the principal stresses (a) σ_1, (b) σ_2, (c) σ_3, in a plane containing the axis of contact between sphere and flat specimen (SS). The diameter of contact is AA, and p_0 is the unit of stress.

by discussing stresses *at* the surface and *below* the surface of the specimen in turn.

Figure 1 indicates the variation of σ_1, σ_2, σ_3 with the radial distance in the specimen surface measured from the center of contact. Figure 3 likewise indicates the stress trajectories in the specimen surface. Within AA all three principal stresses are largely compressive and similar in magnitude; the resulting large component of hydrostatic pressure within this region appears to preclude the possibility of Hertzian fractures initiating there. Outside AA the radially directed stress σ_1 becomes tensile, reaching a maximum value $p_0/6$ at the circle of contact, and thereafter falls off slowly with radial distance. It is this component of the stress that is mainly responsible for the initiation of the Hertzian cracks. The "hoop-stress" σ_2 has a value equal and opposite to that of σ_1 outside AA, and the stress σ_3 normal to the specimen free surface drops to zero. Thus the stress state is one of pure shear outside the circle of contact.

Below the specimen surface the stress distribution

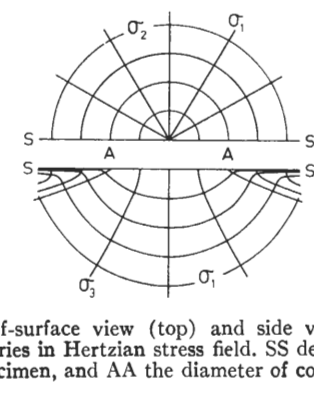


FIG. 3. Half-surface view (top) and side view (bottom) of stress trajectories in Hertzian stress field. SS denotes the surface of the flat specimen, and AA the diameter of contact.

in a plane through the axis of symmetry is indicated by contour maps (Fig. 2) of the three principal stresses. σ_1 is seen to be tensile outside a drop-shaped compressive zone below AA. The magnitude of σ_1 achieves its larger (positive) values only in a shallow region just below AS. The hoop-stress σ_2 becomes tensile at a distance of about $1.7a$ directly below the center of contact, as seen in Fig. 2(b); σ_2 never exceeds σ_1, although it becomes equal to it along the axis of symmetry. The third principal stress, σ_3, never becomes tensile, and at distances from the center of contact large compared with a, it becomes very nearly radially directed (Fig. 3), except very near to the specimen surface, where it rapidly changes its direction to meet the specimen surface orthogonally. Thus, in the region through which the Hertzian cracks are observed to propagate, namely, in the region outside the drop-shaped compressive zone in Fig. 2(a), the two principal stresses σ_1 and σ_3 contained in the plane of symmetry have relative values which depend chiefly on whether they lie in one or the other of two domains: (i) In a very shallow region just beneath the free surface AS the magnitude of σ_1 exceeds that of σ_3, σ_3 being zero or close to zero there; (ii) Below this shallow region, σ_1 drops off rapidly in magnitude, and σ_3 rises in (negative) value, such that the magnitude of σ_3 greatly exceeds that of σ_1.

Crack Extension through Inhomogeneous Stress Fields in Single Crystals

To predict crack paths through an inhomogeneous stress field in any material, it is necessary to consider the stress state at the crack tip. A survey of the literature reveals that the precise crack-tip configuration in brittle solids is not well understood, however, despite many sophisticated attempts to describe it. Griffith[8] was the first to propose a useful crack model; he considered the crack faces to have an elliptical contour, and thereby made use of the stress analysis of Inglis[9] for elliptical cavities. However, calculation indicates that the radius of curvature at the tip of such an elliptical contour would require to be of atomic dimensions if the crack-tip stresses are to exceed the intrinsic cohesive strength of a brittle solid. This, of course, renders any description of the crack configuration close to the tip meaningless in terms of continuum models. Other treatments[10] consider the crack to be representable by a straight cut in the specimen. The main objection to this model is the singularity in the stress field which arises at the crack tip in the continuum elasticity solutions, although it has been suggested that this singularity may be overcome by accounting for nonlinear effects in the elastic behavior near the fracture point. A physically more realistic model was proposed by Elliot,[11] in which two semi-infinite blocks, initially one interatomic distance apart and held together by the attractive forces of cohesion, are allowed to relax into an equilibrium crack configuration under normal tensile loading. Using continuum elasticity theory Elliot computed the geometry of the crack faces and found the crack to close smoothly at its tip. Thus, according to this model, there are atomic bonds at every stage of elongation near the crack tip, so that no stress singularity arises there. To Elliot's procedure Barenblatt[12] raised the valid objection that calculations based on continuum elasticity theory are not applicable to crack faces whose separation is of order one atomic distance. Accepting the essential features of Elliot's model, Barenblatt proposed certain modifications to the description. He argued that a detailed consideration of the cohesive forces acting on the crack faces near the tip requires the crack to close more sharply into a smooth cusp over a characteristic length d: the stresses over the distance d become large, but the smooth closure ensures that the stresses attained are limited by the cohesive strength of the solid. However, a quantitative evaluation of the distance d by Cribb and Tomkins[13] indicates it to be of the order of an interatomic distance. Thus, it would appear that the main obstacle to a satisfactory description of the stress situation at the crack tip lies in the fact that continuum theories are incapable of accurately specifying details of a phenomenon essentially occurring on an atomic scale. Calculations based on lattice models might prove to be a more profitable approach to this problem.

The lack of an exact description of the crack tip does not appear to impose a severe restriction on the ability to predict crack paths in *isotropic* brittle solids. For the diverse approaches of various workers,[14,15,1] although differing in the details of their predictions, lead to the one common conclusion; the ultimate crack path in an inhomogeneously stressed isotropic material tends to become aligned orthogonally to the most tensile of the principal stresses. The various approaches involve the choice of different crack tip models, and the application of different crack extension criteria to determine the direction of incremental crack growth. With regard to the crack extension criteria there are two (essentially equivalent) forms in which these may be stated: (i) The crack extends in that direction from its tip for which the local stress level first exceeds the cohesive strength; (ii) The crack extends in that direction from its tip for which the energy release rate $d(\Delta U)/|\,d\mathbf{c}\,|$ first exceeds 2γ (ΔU is the reduction in mechanical energy due to

[8] A. A. Griffith, Phil. Trans. **A221**, 163 (1920).
[9] C. E. Inglis, Trans. Inst. Naval Archit. **55**, 219 (1913).
[10] M. L. Williams, J. Appl. Mech. **24**, 111 (1957).
[11] H. A. Elliot, Proc. Phys. Soc. **59**, 208 (1947).
[12] G. I. Barenblatt, Advan. Appl. Mech. **7**, 55 (1962).
[13] J. L. Cribb and B. Tomkins, J. Mech. Phys. Solids **15**, 135 (1967).
[14] F. Erdogan and G. C. Sih, J. Basic Eng. **85**, 519 (1963).
[15] B. Cotterell, Internat. J. Fracture Mech. **1**, 96 (1965).

formation of the crack, $d\mathbf{c}$ is an incremental area of the crack, and γ is the specific surface energy corresponding to the newly created surface $d\mathbf{c}$). In an isotropic solid, for which γ (or the cohesive strength) is invariant with orientation, it is only the *maximum* in the energy release rate (or local stress level) which determines the direction of crack extension: in all crack models this maximum tends to direct the extension along a surface delineated by the stress trajectories of the two weaker principal stresses.

In an *anisotropic* material, however, γ is orientation dependent and the direction of crack extension will depend strongly on the *relative anisotropies* in the quantities $d(\Delta U)/|\,d\mathbf{c}\,|$ and 2γ. The degree of angular variation of the energy release rate (or local stress level) with extension direction from the crack tip is not great,[16] but it is dependent on the crack model adopted. For instance, for a slit-like crack, Cotterell[15] calculates the variation in the energy release rate to be less than 10% over an angular range of $\pm 20°$ about the crack tip. The orientation dependence of γ in diamond-structure single crystals is also modest. If the surface energy is calculated as half the work required to break bonds across a given crystallographic plane (hkl),[17] the relation

$$\gamma_{hkl}=[3h^2/(h^2+k^2+l^2)]^{1/2}\gamma_{111} \qquad (4)$$

follows, with $0\leq k,\ l\leq h$. The ratio $\gamma_{hkl}/\gamma_{111}$ is listed for several crystallographic planes in Table I; its value varies between unity [(111) plane] and $\sqrt{3}$ [(100) plane]. Experimental evidence appears to corroborate Eq. (4). For example, the frequency of occurrence of various crystallographic planes in small-scale cleavages[18,5] shows a tendency to decrease with increasing $\gamma_{hkl}/\gamma_{111}$. Also, (4) is consistent with the observation that diamond, silicon, and germanium tend, under conditions of equilibrium crystal growth, to assume an octahedral growth habit.[19] However, Eq. (4) is not exact,[20] so that, together with the uncertainty in the variation of the energy release rate, the application of the crack extension criteria to the determination of crack paths in diamond-type crystals might be expected to give unreliable results. Computer determinations of the crack paths in the Hertzian stress field indeed indicate this to be the case, the computed paths in the diamond crystals depending sensitively on the crack model adopted; by contrast, similarly computed paths for isotropic solids confirm that the crack always tends to follow the σ_2–σ_3 stress trajectory surface, regardless of the crack model. In view of the quantitative inconsistency in the single-crystal case, the observed crack-path features will be restricted to a qualitative interpretation only.

TABLE I. Values of $\gamma_{hkl}/\gamma_{111}$ for some crystallographic planes in crystals with the diamond structure.

Plane	(111)	(221)	(110)	(211)	(210)	(100)
$\gamma_{hkl}/\gamma_{111}$	1.000	1.154	1.225	1.415	1.550	1.732

The hypothesis that crack paths closely follow stress trajectories in isotropic brittle solids is well substantiated by the geometry of cone cracks in glass.[1,2] It is of interest to consider the behavior of these cone cracks in the regions where the stress trajectories curve away from the previous crack plane. If the crack proceeds along its previous plane it must, since it then deviates from the stress trajectories, experience a small component of shear loading. Thus, in order to explain the observed crack behavior, it may be implied that such shear stresses act as restoring stresses whenever a crack deviates from orthogonality to the greatest principal tensile stress. This implication is supported by the evidence of Erdogan and Sih,[14] who subjected pre-existing straight cuts in Plexiglas plates to plane loading: they found that the effect of shear on the cuts was to direct the resulting crack extension from the tips toward that plane which experienced the greatest tensile load. In terms of the crack-tip extension criteria, an equivalent interpretation is that an increasing component of shear loading tends to deflect the direction of maximum energy release rate further away from the existing crack plane toward that plane which experiences the greatest tension. We may now usefully apply this description to the case of anisotropic materials. This time we consider a crack proceeding along a cleavage plane of minimum γ perpendicular to a tensile load. If the crack enters a differently stressed region such that the stress trajectories begin to curve away from the previous crack plane, the question arises as to which of the two factors, the tendency for the crack to follow planes of minimum γ or the tendency for it to follow stress trajectories, has the greater influence on the subsequent crack path. If the crack continues on its original plane, it will experience a small component of shear stress, as before, and this will act to restore the crack to the direction of the stress trajectories. However, if γ has a sufficiently large minimum this tendency may not initially be strong enough, and the crack may thereby proceed along its original cleavage plane. In so doing, the crack will experience an increasing component of shear loading as it progresses, and the restoring tendency will thus become correspondingly stronger. If the anisotropy in γ is not too great this increasing tendency

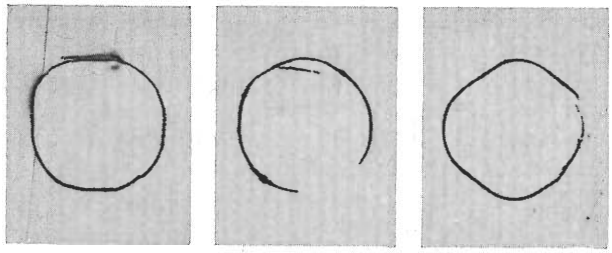
[16] It is difficult to estimate the angular variation of the energy release rate with any accuracy, particularly for large deviations from the straight-ahead direction.

[17] W. D. Harkins, J. Chem. Phys. **10**, 268 (1942).

[18] S. Ramaseshan, Proc. Indian Acad. Sci. **A24**, 114 (1946).

[19] F. C. Frank, *Metal Surfaces; Structure, Energetics and Kinetics* (ASM publication, 1963), Chap. 1.

[20] It is pointed out in the papers by Wolff and Broder and Frank that (4) can only be an approximation, since this equation does not take into account the effects of temperature or relaxation of surface atoms on the value of γ.



(a) (b) (c)

FIG. 4. Traces of Hertzian cracks on (a) (100), (b) (111), (c) (110) surfaces of silicon. Specimens lightly abraded, indented, etched, and viewed in normally reflected light. $r=0.64$ cm, $a=0.05$ cm. (Some abrasion scratches are still visible. Incompleted traces are due to small deviations from perpendicularity between specimen surface and line of application of load.)

will ultimately deflect the crack. Thus, depending on the degree of anisotropy of the material, and the nature of the stress field through which the crack propagates, we might, in general, expect the crack to compromise between the tendencies to follow stress trajectories and planes of minimum γ.

Formation of the Hertzian Cracks in Diamond-Structure Materials

One of the most striking features of a Hertzian pressure crack in diamond,[21–23] germanium,[24,25] or silicon (Fig. 4) is the near-polygonal surface trace which reflects the symmetry of the diamond structure. (Analogous behavior is also observed in crystals with the zincblende structure,[26] although, because of their partially ionic bonding character, these crystals have different cleavage properties.[5]) From an analysis of such traces the initial stage of crack growth near the crystal surface appears to proceed along favorably oriented {111} cleavage planes. However, on examination of the crack below the crystal surface it is always found that the path tends to deviate away from {111} planes toward the Hertzian cone. Figure 5, whose plane bisects diametrically opposite sides of the (111) surface trace in Fig. 4(b), illustrates this feature. The double arrows in Fig. 5 indicate traces of the {111} cleavage planes, and it is seen that both sides of the internal crack spread downward and outward and away from these planes. (The fact that the vertical plane perpendicular to the plane of Fig. 5 is not one of crystallographic symmetry makes itself evident in the geometry of the internal crack.) In a similar manner, the internal cracks deviate outward from the {111} planes extending into the crystal from the sides of the surface traces on the (100) and (110) surfaces. Thus, there appears to be a certain predisposition for the crack to propagate along cleavage planes near the crystal surface, and for it to follow more closely the stress trajectories of the two lesser principal stresses as it propagates into the interior.

In examining the growth of a Hertzian fracture in glass, it was concluded that in most instances the crack initiates at or very close to the circle of contact and runs around this circle (which corresponds to a σ_2 stress trajectory), but with a somewhat exaggerated radius of curvature because of a "pseudo-inertia" effect, producing first a shallow surface ring crack eccentric relative to the circle of contact.[1] The hypothesis that the surface ring forms before subsequent downward extension into the material was based on the knowledge that the Hertzian stresses acting across the ultimate crack path barely decrease from their maximum value at the starting point as one follows the surface trace and that, in marked contrast, they drop off extremely rapidly with distance measured along the downward crack path. Although in the present case of interest the surface traces deviate further from the circle of contact than their counterparts in glass, the weight of the above argument is barely lessened. One may therefore infer that the crack initiates at a point close to the circle of contact, runs around both sides in an unstable manner to complete itself on the opposite side, and subsequently, on application of a further load on the spherical indenter, propagates downward. This inference has some experimental support,[21] and calculations in a later section serve to substantiate it.

This model of crack formation permits one to treat the problem of Hertzian crack formation in two convenient stages. First, in treating the surface ring crack we need only consider that shallow stress domain near the free surface where the tensile stresses have their larger values. If, by virtue of its crystallographic cleavage tendency, the surface crack deviates from the circle of contact as it propagates around the indenter, it will suffer a shear loading whose magni-

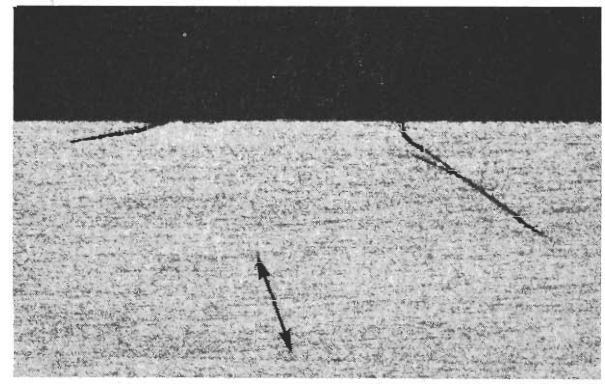


FIG. 5. Cross-sectional profile of Hertzian crack made on the (111) surface in Fig. 4(b). Plane of diagram is $(01\bar{1})$. Double arrow indicates trace of {111} easy-cleavage plane. Specimen sawn through crack, ground, lightly etched, and viewed in obliquely reflected light.

[21] V. R. Howes and S. Tolansky, Proc. Roy. Soc. **A230,** 287, 294 (1955).
[22] S. Tolansky and V. R. Howes, Proc. Phys. Soc. **B70,** 521 (1957).
[23] B. R. Lawn and H. Komatsu, Phil. Mag. **14,** 689 (1966).
[24] E. N. Pugh and L. E. Samuels, Phil Mag. **8,** 301 (1963).
[25] O. W. Johnson, J. Appl. Phys. **37,** 2521 (1966).
[26] J. W. Allen, Phil. Mag. **4,** 1046 (1959).

tude will depend on the relative values of the principal stresses. Thus, when the surface crack has an angular deviation α from a σ_2 stress trajectory the resulting shear stress τ_{12} across the crack path is related to the tensile stress σ_1 (the intermediate stress $\sigma_3 \simeq 0$ playing only a secondary role in the fracture process) by

$$\tau_{12} = (\sigma_1 - \sigma_2) \sin\alpha \cos\alpha.$$

For small angular deviations we may write

$$\tau_{12}/\sigma_1 \simeq [1 - (\sigma_2/\sigma_1)]\alpha.$$

Since $\sigma_2 \simeq -\sigma_1$ near the specimen surface outside the area of contact between indenter and specimen, this reduces to

$$\tau_{12}/\sigma_1 \simeq 2\alpha. \tag{5a}$$

Similarly, with regard to the subsequent downward extension of the crack, we need to take into account that second stress domain beyond the shallow surface region, where σ_1 becomes small (but remains tensile). In this case, the crack tends to follow the σ_3 stress trajectories (σ_2 becoming the intermediate stress), and the corresponding shear stress τ_{13} across the crack path due to a small angular deviation α from the trajectory path is given by

$$\tau_{13}/\sigma_1 \simeq [1 - (\sigma_3/\sigma_1)]\alpha.$$

Within this second stress domain, $\sigma_3 \simeq -10\sigma_1$, so that

$$\tau_{13}/\sigma_1 \simeq 11\alpha. \tag{5b}$$

Thus, comparing (5a) and (5b), we see that small deviations from the stress trajectories give rise to considerably larger restoring shear stresses once the crack leaves the shallow surface region. Hence we would expect the anisotropy in γ to influence the crack path more strongly near the crystal surface, where the restoring effect is least. This would explain the crystallographic cleavage characteristics of the surface trace, and the resemblance between the profiles of the downward extending crack and the conventional cone crack in glass.

In the above treatment, the effects of "pseudo-inertia" were mentioned, but not discussed in detail. Its presence will lead only to a small eccentricity of the crack geometry with respect to the axis of symmetry, and for present purposes may be neglected. A more important effect may be a possible mutual interference between adjacent sections of the crack propagating downward from the sides of the near-polygonal surface traces. For example, alternate sides of the near-hexagonal surface trace on the (111) surface in Fig. 4(b) will extend downward along geometrically different paths, as shown in Fig. 5; this will presumably give rise to a difficulty in accommodating the different geometries in the regions common to the adjacent sides. This effect may explain the observation of fracture irregularities at the internal fracture interfaces in diamond by Howes and Tolansky.[21,22] However, as will be shown in the next section, the mechanics of the downward extending crack are critically determined by its behavior down to a depth of about $a/10$, where such secondary effects should have little influence.

Mechanics of Fracture

A. Theoretical

Following the procedure adopted by Frank and Lawn,[1] we calculate the *energy release rate* $\mathcal{G} \equiv d(\Delta U)/|\, d\mathbf{c}\,|$ per unit width of the downward extending Hertzian crack front. For a plane crack in a semi-infinite medium, we may write[27]

$$\mathcal{G} = (\pi/E)(1-\nu^2)(\mathcal{K}_1^2 + \mathcal{K}_2^2), \tag{6}$$

where E = Young's modulus, ν = Poisson's ratio, and the $\mathcal{K}$ terms, the *stress intensity factors*, are given by

$$\mathcal{K}_1 = \frac{2}{\pi} c^{1/2} \int_0^c \frac{\sigma(b)\, db}{(c^2 - b^2)^{1/2}}$$

$$\mathcal{K}_2 = \frac{2}{\pi} c^{1/2} \int_0^c \frac{\tau(b)\, db}{(c^2 - b^2)^{1/2}}, \tag{7}$$

c being the crack length, b the distance measured along the crack path, and $\sigma(b)$ and $\tau(b)$ the *prior* normal and shear stresses along the crack path. We now make use of the geometrical similarity of the Hertz elastic problem by expressing stresses in terms of p_0 and lengths in terms of a; thus, it becomes convenient to redefine (7) in terms of the dimensionless quantities

$$K_1 = \frac{2}{\pi}\left(\frac{c}{a}\right)^{1/2} \int_0^{c/a} \frac{(\sigma/p_0)\, d(b/a)}{(c^2/a^2 - b^2/a^2)^{1/2}}$$

$$K_2 = \frac{2}{\pi}\left(\frac{c}{a}\right)^{1/2} \int_0^{c/a} \frac{(\tau/p_0)\, d(b/a)}{(c^2/a^2 - b^2/a^2)^{1/2}}, \tag{8}$$

so that (6) becomes

$$\mathcal{G} = (\pi/E)(1-\nu^2)(K_1^2 + K_2^2)\, p_0^2 a. \tag{9}$$

We can eliminate p_0 and a, and express (9) in terms of the more readily observable parameters P and r, using (1) and (3); thus,

$$\mathcal{G} = 3P(1-\nu^2)(K_1^2 + K_2^2)/4\pi kr. \tag{10}$$

Now the Griffith[8] energy-balance criterion states that a crack will extend when

$$\mathcal{G} \geq 2\gamma. \tag{11}$$

Combining (11) with (10) and (4), the condition for crack extension in diamond structure crystals becomes

$$3P(1-\nu^2)(K_1^2 + K_2^2)/4\pi kr \geq 2[3h^2/(h^2+k^2+l^2)]^{1/2}\gamma_{111}. \tag{12}$$

[27] G. R. Irwin *Handbuch der Physik* **6**, 551 (Springer-Verlag, Berlin; 1958).

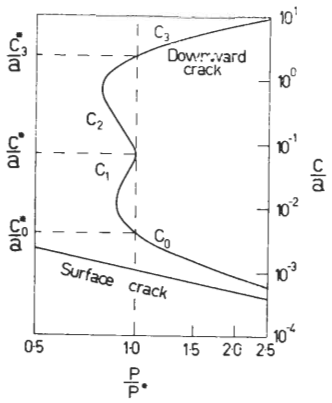


FIG. 6. Plot of c/a as function of P/P^* for downward-extending crack (upper curve) and surface crack (lower curve) within the Hertzian stress field in an isotropic elastic medium. (For description, see text.)

The quantities $(K_1^2+K_2^2)$ and $[3h^2/(h^2+k^2+l^2)]^{1/2}$ may be computed numerically at any point along the crack path, once the crack configuration is known, so by collecting these into one term

$$\phi(c/a)=[3h^2/(h^2+k^2+l^2)]^{1/2}/(K_1^2+K_2^2) \quad (13)$$

we may rearrange (12) to obtain

$$P(c/a)\geq[8\pi kr\gamma_{111}/3(1-\nu^2)]\phi(c/a) \quad (14)$$

as the condition that the crack should extend.

In applying (14) to the downward-extending crack in diamond-structure crystals, we make the implicit assumption that the plane crack formulas (6) and (7) are applicable. Although the Hertzian crack is curved, this assumption will not lead to large errors for $c\ll a$; as we shall see below the fracture process is critically determined in this early stage of crack growth. A more serious source of inaccuracy arises from the inability to compute crack paths in single crystals in a quantitative manner. However, for various surface orientations of the diamond crystals the possible crack paths, either (i) observed experimentally by sectioning crystals (e.g., Fig. 5) or (ii) computed using different crack models, all lie outside the drop-shaped compressive zone [Fig. 2(a)] and thereby give rise to a function $\phi(c/a)$ with a maximum at $c\simeq a/10$; it is the *existence*, and not the *value*, of the maximum in $\phi(c/a)$, which turns out to be responsible for the form of Auerbach's law, inaccuracies in the treatment affecting only the numerical values of the constant in this law. In all cases, the computed $\phi(c/a)$ curves[28] differ little in form from the more accurately specifiable curve corresponding to the cone crack in glass. The latter curve is represented in Fig. 6 by a plot of the ratio P/P^* as a function of the relative *equilibrium* crack length c/a (the asterisk denoting a value corresponding to the maximum in the $\phi(c/a)$ curve, so that for an equilibrium crack (14) reduces to $P/P^*=\phi/\phi^*$). The different branches of this curve, distinguished by the labels c_0, c_1, c_2 and c_3, represent different types of crack equilibria, whose significance has been discussed elsewhere.[1] It will be sufficient to state here that the equilibrium is stable or unstable according to whether or not dP/dc is positive or negative. (For a general discussion on crack stability, reference is made to the paper by Barenblatt.[12]) Also included in Fig. 6 is the curve corresponding to a crack extending from a point on the circle of contact along a path subjected to an undiminishing stress. Such a curve closely represents the conditions under which the surface ring crack propagates around the indenter. Thus, the two curves in Fig. 6 may be regarded as indicative of the two stages of formation of the Hertzian crack in the diamond-type crystals. Since (14) is satisfied for all points lying to the right of the curves in the figure we conclude that the surface crack should extend at loads lower than those required to extend the downward crack.

We now consider the initial stage of crack growth to begin from a submicroscopic flaw, of characteristic length c_f, located close to the circle of contact. As we increase the load P on the indenter, the point $(P/P^*,\ c_f/a)$ will intersect first the lower curve, thereby satisfying (14), and the surface ring will then run around the indenter in an unstable manner to intersect itself on the opposite side. With a further increase in load the point $(P/P^*,\ c_f/a)$ will then intersect the upper curve; for larger values of c_f it is seen that an increasingly larger load increment will be required to propagate the downward extending crack through the rapidly diminishing stress field beneath the specimen surface after the surface crack has been formed. The intersection of major interest here occurs when the flaw size lies within the range $c_0^*\leq c_f\leq c^*$. If the point $(P/P^*,\ c_f/a)$ intersects the c_0 branch an incremental increase in load will cause the crack to proceed unstably toward the c_1 branch at constant load. When the stable c_1 branch is reached, or if intersection first occurs there, further gradual increases in load will be required to make the crack grow along the curve in a stable manner until $c>c^*\simeq a/10$ (or, $P>P^*$). This final stage in which the crack is "pushed over the hill" to proceed unstably at constant load to $c\sim 3a$ represents the development of the visible Hertzian crack. From (14) the condition for this to happen is

$$P(c/a)\geq[8\pi kr\gamma_{111}/3(1-\nu^2)]\phi(c^*/a).$$

The value of the constant $\phi(c^*/a)$ calculated from (13) is of the order 4×10^4 for the various possible crack paths. Thus, the condition simplifies to

$$P^*/r\simeq 4\times10^5 k\gamma_{111}, \quad (15)$$

which establishes Auerbach's law and specifies that

[28] It should perhaps be remarked that the procedure adopted here of numerically evaluating $\phi(c)$ along the crack path replaces the more analytical (but less accurate, because of the many approximations necessary) treatment presented in Ref. 1 for establishing the functional form of $P(c)$.

the Auerbach constant is proportional to the specific surface energy. We note here that (15) is independent of the flaw size c_f. As long as c_f remains within the range $c_0^* \leq c_f \leq c^*$ the critical load required to make the Hertzian crack extend to c_3 will always be equal to P^*. For flaws outside this size range, however, Auerbach's law will break down, and flaw statistics will largely determine the critical load.[1]

B. Experimental

Some indentation tests were performed on silicon surfaces with a view to investigating the Auerbach law. All specimens were cut to a thickness of about 1 cm from one single crystal,[29] with test surfaces oriented to within 1° of (100), (110), or (111) planes. The tests were carried out on an Instron testing machine, using sintered tungsten carbide Brinell balls as indenters; the crosshead was driven at a speed of 0.005 cm·sec^{-1} in all experiments described below. Because of the difficulty in determining the precise instant of fracture in the opaque silicon specimens a systematic series of indentations, with maximum load as variable, was made on each specimen with five balls. The specimens were then etched in Dash's etchant,[30] thereby revealing which indentations had produced fully developed Hertzian fractures. The critical fracture load could, under favorable conditions of specimen preparation (see below), be determined to within about ±5% by this method, an uncertainty which accounts for nearly all the scatter in the results presented below.

The first tests were made on specimens prepared carefully by using 0.3 μm alumina powder as a final mechanical polish and by following with a light chemical polish.[30] The measured critical load for a given indenter on these specimens showed a scatter of up to 100%. This was attributed to a lack of sizable flaws within the range $c_0^* \leq c_f \leq c^*$ necessary for Auerbach's law to hold. The specimen surfaces were then lightly abraded with No. 600 SiC paper; according to the study by Stickler and Booker[31] such treatment introduces an abundance of surface cracks down to a depth of about 10 μm. This procedure, by increasing the flaw *density* as well as the flaw *size*, improves the likelihood of initiating the Hertzian crack at the most favorable location on the circle of contact. Putting $c_f \simeq 10^{-3}$ cm, and taking $c_0^* \simeq 5\times10^{-3}a$ and $c^* \simeq 10^{-1}a$ from Fig. 6, the condition that Auerbach's law should hold now becomes approximately $5\times10^{-3}a \leq 10^{-3}$ cm $\leq 10^{-1}a$. By eliminating P from (15) and (1), and substituting the resulting expression for a into this condition, we arrive at upper and lower limits of indenter radius within which Auerbach's law might be expected to hold: inserting $\gamma_{111} \simeq 1.4\times10^3$ erg cm^{-2} and $E \simeq 1.3\times10^{12}$ dyn cm^{-2} for silicon, and $k \simeq 0.8$ for a tungsten carbide indenter on a silicon specimen, we thus have, approximately, 5×10^{-2} cm $\leq r \leq 5$ cm. This embraces the range of indenter size used in these experiments. The results obtained with the abraded specimens are shown in Fig. 7 in which P^*/r is plotted as a function of r for the three surface orientations used. Within the experimental scatter it is seen that Auerbach's law holds in each case (the full lines drawn through the points representing mean values of P^*/r).

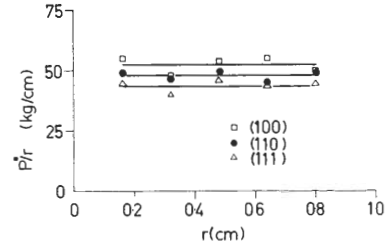


FIG. 7. P^*/r as function of r for (100), (111), (110) surface orientations of silicon.

A second series of tests was performed in order to reveal any possible effect of a variant flaw size c_f on Auerbach's constant. We saw earlier that there should be no such effect on P^*/r, as long as $c_0^* \leq c_f \leq c^*$. Indentations were made on specimens abraded with No. 320 SiC paper which, according to Stickler and Booker, should introduce surface cracks down to about 20 μm. The results showed no systematic deviation from the data in Fig. 6. This serves to confirm that the Auerbach constant is independent of flaw size, so that no stringent surface preparation of specimen samples, other than ensuring that an adequate amount of flaws are present, should be necessary in the Hertzian test.

DISCUSSION

In the foregoing sections we discussed the formation of the Hertzian crack as the load on the indenter was increased, and concluded that the fully developed crack would have a length of several times the radius of the circle of contact. It is now of interest to consider the likely behavior of the crack upon release of the load. If the process were to be completely reversible the crack faces would "heal" perfectly. Of course, *complete* reversibility can never be realized in practice, although in some carefully controlled cleavage experiments on mica in high vacuum the crack interface can be opened and closed repeatedly with little energy loss.[32] In other crystals with less perfect

[29] "Lopex" grade silicon, n-type, specific resistance 1.5 Ω cm^{-1}, dislocation density <500 cm^{-2}, purchased from Texas Instruments.

[30] W. C. Dash, J. Appl. Phys. **29**, 705 (1958). The Dash etchant 1 HF, 3 HNO_3, 8–12 acetic acid (parts by volume) preferentially attacks regions of strain around crystalline defects. The same etchant, but with the acetic acid omitted, serves as a chemical polish.

[31] R. Stickler and G. R. Booker, Phil. Mag. **8**, 859 (1963).

[32] For a summary of cleavage experiments on mica see F. P. Bowden and D. Tabor, *The Friction and Lubrication of Solids*, (Clarendon Press, Oxford, 1964), Pt. II. Chap. 20.

cleavages, such as LiF,[33] partial healing can be observed to occur at cleavage interfaces, with certain factors such as cleavage step formation, production of dislocation arrays, etc., preventing opposite faces from closing in complete registry. As far as the Hertzian cracks are concerned, there is good evidence from optical observations in transparent specimens such as glass and diamond[23] that the internal crack does retreat upon removal of the indenter load, but that cleavage steps and other fracture disturbances at the crack interface prevent complete closure. The case of diamond is of special note, since it implies that silicon and germanium, which exhibit very similar fracture characteristics, should behave in the same way.

It is with the opaque semiconductor crystals that much controversy has arisen in the past. Allen,[26] in discussing the damage produced by dropping steel spheres on to InSb surfaces, postulated that the cracks thus formed closed imperfectly to form an interface containing a high density of dislocations and other defects. This mismatch interface Allen termed a "dislocation crack." Pugh and Samuels,[24] however, in examining similar damage in germanium, objected to Allen's mechanism on the grounds that the same residual damage was observed after static and impact loading; they claimed that different amounts of contaminating gas should penetrate the open cracks and thus give rise to different healing behaviors. On the basis of preferential etching experiments, Pugh and Samuels suggested, as an alternative mechanism, that dislocations be *nucleated*, but not *propagated*, in the intense local stress field under the indenter, so that the residual dislocation crack could be explained without invoking a healing process at all. On the other hand, Johnson,[25] in further work on the indentation of germanium, ruled out the possibility of dislocation glide processes accompanying fracture, and also pointed out reasons why the rate of entrance of contaminating gases into the crack might not significantly affect a healing process. His conclusions led him to favor the Allen process. It is, perhaps, surprising that the workers on the semiconductor materials appear to have been unaware of the parallel behavior in diamond, in which the cracks are visually observed to open and close as the indenter is depressed and withdrawn.[21,22] This evidence, together with the observations of the healing characteristics of other crystals in cleavage experiments,[34] lends strong support to the Allen concept. In terms of the theoretical curve for the downward crack in Fig. 6 we would therefore predict the fully developed Hertzian crack to decrease in length along an extrapolation of the c_3 branch through the origin as the load on the indenter is decreased from its critical value to zero, the presence of fracture irregularities at the crack interface severely inhibiting the healing process in localized regions.

Experimental and theoretical considerations indicate that Auerbach's law should hold for the diamond structure crystals, or, indeed, for any crystals where anisotropies in the elastic constants and the surface energy are moderate. Therefore, because the approach to the Auerbach condition, $c \geq c^*$, is necessarily one of stable crack growth along the c_1 branch in Fig. 6, the measurement of the Auerbach constant P^*/r in (15) affords a convenient means of gaining information about the reversible fracture surface energy γ in such crystals. Of course, due to the approximations made in the theoretical treatment any such determination of the *absolute* value of γ will be unreliable, but *relative* values may, in principle, be measured with the same degree of accuracy as the Auerbach constant. This raises the possibility of using the Hertzian test to investigate certain mechanical properties of brittle crystals. For instance, the fracture surface energy (and hence the inherent strength) may be measured as a function of temperature (provided no significant change in the crystallographic anisotropy occurs), environmental atmosphere, irradiation dosage, impurity content, loading conditions (e.g., loading rate, cyclic loading), photon illumination (photomechanical effect), electric current through the specimen (electromechanical effect), etc., simply by recording the variation of the Auerbach constant. And the Hertzian test has the unique advantage that the Auerbach condition is independent of crack size or statistics of distribution, so that small specimens requiring no stringent preparation other than ensuring a reasonably flat, flaw-abundant surface, may be used.

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Professor F. C. Frank, University of Bristol, England, for many discussions and for his interest in this work. The following also assisted with their useful comments: Dr. S. Burns and Professor J. Rice (Brown University), J. Craig (Defence Standard Laboratories, Sydney, Australia), Dr. J. Cribb (Colonial Sugar Refineries, Sydney, Australia), and R. Khokhar and Professor D. Haneman (these laboratories), who kindly gave permission to mention their results[34] prior to publication. Part of this work was financed by a grant from Industrial Distributors (Sales) Ltd., and part by an Advanced Research Projects Agency contract.

[33] These observations are based on optical and x-ray topography techniques: A. J. Forty, Proc. Roy. Soc. **A242**, 392 (1957); A. J. Forty and C. T. Forwood, Trans. Brit. Ceram. Soc. **62**, 715 (1963); C. T. Forwood and B. R. Lawn, Phil. Mag. **13**, 595 (1966).

[34] In addition to the experiments mentioned earlier in the Discussion related experiments have been performed in this laboratory on silicon by R. Khokhar and D. Haneman. Using the method of transmission x-ray topography the location of the crack tip in a cantilever cleavage specimen has been observed to retreat on partial withdrawal of the cleavage wedge.